

<https://helda.helsinki.fi>

---

## Structural validity of the Finnish Manchester-Oxford Foot Questionnaire (MOXFQ) using the Rasch model

Ponkilainen, Ville T.

2021-01

---

Ponkilainen , V T , Miettinen , M , Sandelin , H , Lindahl , J , Häkkinen , A H , Toom , A , Tillgren , T , Ilves , O , Latvala , A O , Ahonen , K , Sirola , T , Sampo , M , Väistö , O & Repo , J P 2021 , ' Structural validity of the Finnish Manchester-Oxford Foot Questionnaire (MOXFQ) using the Rasch model ' , Foot and Ankle Surgery , vol. 27 , no. 1 , pp. 93-100 . <https://doi.org/10.1016/j.fas.2020.02.012>

---

<http://hdl.handle.net/10138/338155>

<https://doi.org/10.1016/j.fas.2020.02.012>

---

cc\_by\_nc\_nd

draft

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

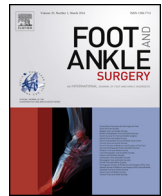
*Please cite the original version.*



Contents lists available at ScienceDirect

Foot and Ankle Surgery

journal homepage: [www.elsevier.com/locate/fas](http://www.elsevier.com/locate/fas)



## Structural validity of the Finnish Manchester-Oxford Foot Questionnaire (MOXFQ) using the Rasch model

Ville T. Ponkilainen<sup>a,\*</sup>, Mikko Miettinen<sup>b</sup>, Henrik Sandelin<sup>b,c</sup>, Jan Lindahl<sup>b</sup>,  
Arja H. Häkkinen<sup>d,e</sup>, Alar Toom<sup>a</sup>, Tomi Tillgren<sup>a</sup>, Outi Ilves<sup>d</sup>, Antti O. Latvala<sup>f</sup>,  
Katri Ahonen<sup>a</sup>, Timo Sirola<sup>b</sup>, Mika Sampo<sup>g</sup>, Olli Väistö<sup>h</sup>, Jussi P. Repo<sup>a</sup>

<sup>a</sup> Department of Surgery, Central Finland Central Hospital, Jyväskylä, Finland

<sup>b</sup> Department of Orthopedics and Traumatology, Helsinki University Hospital and University of Helsinki, Finland

<sup>c</sup> Orthopaedic Department, Liverpool Hospital, Sydney, New South Wales, Australia

<sup>d</sup> Health Sciences, Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

<sup>e</sup> Department of Physical Medicine, Central Finland Health Care District, Jyväskylä, Finland

<sup>f</sup> Department of Orthopedics and Traumatology, Oulu University Hospital, Oulu, Finland

<sup>g</sup> Comprehensive Cancer Center, Helsinki University Hospital and University of Helsinki, Finland

<sup>h</sup> Department of Orthopedics and Traumatology, Seinäjoki Central Hospital, Oulu, Finland

### ARTICLE INFO

#### Article history:

Received 30 September 2019

Received in revised form 16 February 2020

Accepted 22 February 2020

Available online xxx

#### Keywords:

Foot  
Ankle  
Patient-reported  
MOXFQ  
PROM  
Psychometrics  
Validation  
Surgery  
Rasch

### ABSTRACT

**Background:** The 16-item patient-reported Manchester-Oxford Foot Questionnaire (MOXFQ) with subscales of pain, social interactions, and walking/standing has been claimed for strongest scientific evidence in measuring foot and ankle complaints. This study tests the validity of the Finnish MOXFQ for orthopaedic foot and ankle population using the Rasch analysis.

**Methods:** We translated the MOXFQ into Finnish and used that translation in our study. MOXFQ scores were obtained from 183 patients. Response category distribution, item fit, coverage, targeting, item dependency, ability to measure latent trait (unidimensionality), internal consistency (Cronbach's alpha), and person separation index (PSI) were analyzed.

**Results:** Fifteen of the items had ordered response categories and/or sufficient fit statistics. The subscales provided coverage and targeting. Some residual correlation was noted. Removing one item in the pain subscale led to a unidimensional structure. Alphas and PSIs ranged between 0.68–0.90 and 0.67–0.92, respectively.

**Conclusions:** Despite some infractions of the Rasch model, the instrument functioned well. The subscales of the MOXFQ are meaningful for assessing patient-reported complaints and outcomes in orthopaedic foot and ankle population.

© 2020 European Foot and Ankle Society. Published by Elsevier Ltd. All rights reserved.

## 1. Introduction

Clinical outcome assessment is one of the cornerstones of improving the health care system in patient care [1–4]. The patient-reported outcome measures (PROMs) have become widely used in assessing the treatment outcomes in the health care system [5]. Their increased use has also been noted after lower extremity surgery as up to 140 different PROMs have been used for assessing this anatomical region [6,7]. In foot and ankle surgery, the quality of the PROMs has had remarkable variability in

their validity, reliability, responsiveness, and other psychometric properties [7–11]. Due to the wide selection and variability of available PROMs for foot and ankle assessment, testing of the measurement properties and quality assessment of PROMs is needed before making a choice which instrument to use for foot and ankle patients with a specific disease [12–15].

The Manchester-Oxford Foot Questionnaire (MOXFQ) was originally developed for evaluating outcomes after hallux valgus surgery [16]. It has shown to be the most suitable PROM for patients with hallux valgus in terms of psychometric properties [17]. The MOXFQ has also been proven to be valid in foot and ankle surgery [18] and to represent the most evidence for quality in assessing foot and ankle [8]. It has been praised for its high readability and superior psychometric properties compared to several other instruments for assessing foot and ankle pathologies [8,19].

\* Corresponding author at: Department of Surgery, Central Finland Central Hospital, Keskussairaalantie 19, 40620 Jyväskylä, Finland.

E-mail address: [ville.ponkilainen@tuni.fi](mailto:ville.ponkilainen@tuni.fi) (V.T. Ponkilainen).

The MOXFQ has been extensively tested using the classical test theory [16,18,20–27] and other traditional psychometric tests [8]. However, its psychometric properties have never been investigated for the whole extent of foot and ankle patients using the Rasch Measurement Theory (RMT) [28]. The Rasch mathematical modeling tests how the data fits into a predefined model [29]. The model is based on a theory of a latent trait and additive conjoint measurement [30]. The RMT could provide valuable information of the construct of the MOXFQ in assessing pain, walking/standing and social interaction among foot and ankle patients [31–33]. As the MOXFQ has shown to be a good prospect in foot and ankle patient-reported outcome assessment, it is highly important to thoroughly test its construct validity for assessing outcomes foot and ankle pathologies. The Consensus-based Standards for the selection of health status Measurement INstruments (COSMIN) initiative recommends that the structural validity, for example using the Rasch model, should be tested after assessing the content validity of a PROM [34]. The MOXFQ has been tested for its unidimensionality using the factor analysis and Rasch model for hallux valgus [16]. Nonetheless, the MOXFQ has not been tested using the Rasch model for other orthopaedic foot and ankle pathologies. The present analysis provides important information about the suitability of the MOXFQ for the whole spectrum of orthopaedic foot and ankle patients. The aim of the present study is to further validate the Finnish MOXFQ among orthopaedic foot and ankle population using Rasch measurement analytic techniques.

## 2. Materials and methods

Patients were invited to participate in the study from four different orthopaedic centers in Finland as follows: Central Finland Central Hospital, Jyväskylä; Oulu University Hospital, Oulu; Peijas Hospital, Helsinki University Hospital, Vantaa; Seinäjoki Central Hospital, Seinäjoki. Patient inclusion criteria were age 18 years or older, full understanding of Finnish language, planned surgery for foot and/or ankle. Exclusion criteria were unbalanced mental disorder (such as schizophrenia or depression) that could negatively affect the reliability of the testing, age less than 18 years, incomplete ability to understand Finnish, no scheduled foot and/or ankle surgery.

Patients were invited to participate either by mail or face-to-face before the surgery. Participants completed questionnaires and a pre-information form on the day of surgery.

All participants were provided with information about the study and they signed a consent form for agreement to participate according to good ethical principles described in more detail in the Declaration of Helsinki [35]. Ethical review board of Central Finland, Helsinki and Uusimaa, Northern Ostrobothnia, and South Ostrobothnia health care districts approved the study protocol.

### 2.1. Manchester-Oxford Foot Questionnaire (MOXFQ)

The MOXFQ is a patient-reported outcome measure that was developed to evaluate the outcomes after treatment for hallux valgus [16]. However, it has also been shown to be valid for foot and ankle surgery [36,37]. The instrument contains 16 items and the response options are on a five-point Likert-scale from 1 to 5 [16]. The MOXFQ can be divided into three domains: walking/standing (7 items), pain (5 items), and social interaction (4 items) [37]. The scores of the subscales are scaled from 0 to 100, where 0 represents low symptoms and 100 the most severe symptoms [37]. In addition, it can be presented as a summary score [37]. The scales' minimally important change has been determined to be 24.75 points [38]. The questionnaire has previously been translated and validated into Korean [22], Dutch [20], Turkish [23], German [21], Spanish [26], Persian [24], and Italian [25] languages.

### 2.2. Translation and cultural adaptation

The translation took place according to the guidelines provided by the copyright holder (Oxford University Innovation Limited, Oxford, UK). The process is described in more detail below.

Two translators who have Finnish as their native language and are fluent in English produced two forward translations independently. Both translators have over four years of experience for translation of PROMs and have expertise in medical translations. A reconciliation of the two versions and a written report was produced by the two translators. A back-translation into English and written reports were produced by two translators independently. The translators have English as their native language and are fluent in Finnish and familiar with Finnish culture. The two translators were not familiar with the contents of the original MOXFQ instrument.

The project manager reviewed the back translations against the source version. The pre-test version was formatted to match the original layout in the English questionnaire. The pre-test version was thereafter pilot-tested among five foot and ankle patients and cognitively debriefed by the in-country investigator to reveal any discrepancies in the translated version. Pre-testing using five patients also adheres to the guidelines by Wild et al. [39]. A written report was produced. The project manager reviewed the results of the pilot-testing. Minor semantic changes in the structure of the questionnaire were made in items 9 "I feel self-conscious about my foot" and 10 "I feel self-conscious about the shoes I have to wear" for accurately capture what is relevant to the patient.

A multidisciplinary panel reviewed the outcomes and a pre-final version was introduced. The pre-final version was proofread by the in-country investigator and a language expert. Refinements were made in descriptions section on the first page and in item 14 "I am unable to do all my social or recreational activities because of pain in my foot" to amend the linguistics. We believe that the consistency of language use is likely to be good in a small population (Finland). However, when adapting this kind on translation process, it should be kept in mind that similar process might not work out well in larger population with multiple variations of the same language.

### 2.3. Statistics

Sociodemographic and clinical data are reported as means with standard deviation (SD), ranges, or numbers with percentages.

The RMT was applied to the data to investigate the psychometric properties of the MOXFQ. The RMT serves for construct validation. These analyses were based on statistical and illustrative tests. Andrich Rating Scale Model was used. The approaches of RMT are extensively explained in more detail elsewhere [28,31–33,40–46]. Most pertinent analyses and tests used in this study are described below.

#### 2.3.1. Item threshold distribution

The response categories were investigated to test how they would perform in this patient sample. In this contest, a threshold refers to the distinct point where the patient has a 50% possibility to choose one of the two response options. An optimal response category distribution would follow the Gaussian curve [47]. In foot and ankle PROMs, fusing one or more response categories could potentially have positive impact on their measurement accuracy [48].

#### 2.3.2. Fit statistics

The authors used three approaches to investigate fit statistics. Item-person interaction (log residuals), item-trait interaction (chi-square values) and item characteristic curves were applied. Fit

**Table 1**  
Background characteristics of the patients.

	N = 183
Age, mean (SD)	53 (15)
Female, n (%)	127 (69)
BMI, mean (SD)	27.6 (6.0)
Duration of foot or ankle complaints, years, median (IQR)	8 (2–12)
Affected foot or ankle previously operated, n (%)	79 (43)
Indication for operative treatment, n (%)	
Deformity of foot or ankle	75 (41)
Osteoarthritis of foot or ankle	47 (26)
Flat foot or cavoid foot	13 (7)
Ankle instability	2 (1)
Other	46 (25)
General health state (Likert), mean (SD)	2.9 (0.7)
Education, n (%)	
Higher education	75 (41)
Upper secondary level education	21 (11)
Basic education	75 (41)

residuals values between  $\pm 0.25$  were accepted for item fit. Fit values exceeding this threshold were considered as potentially misfitting. Non-significant P-values were hypothesized to be observed after Bonferroni adjustment.

### 2.3.3. Coverage and targeting

Item-person distribution was investigated to reveal how the MOXFQ provides coverage and targeting for the study population. Non-significant P-values between MTPJ1 pathologies and foot and ankle pathologies were hypothesized to be noted after Bonferroni adjustment.

### 2.3.4. Item dependency

Residual correlation between each two items was estimated to reveal whether there would be any dependency between the items. Dependency could potentially refer to lower reliability and redundancy of items. Threshold for residual correlation was set at  $\geq 0.3$  based on accepted general assumption.

### 2.3.5. Unidimensionality

The scales were tested for unidimensionality to investigate whether the scales of the MOXFQ would measure a single trait (pain, mobility or social interaction). Item residual loadings of  $\pm 0.3$  were used to build subsets for testing of unidimensionality. A threshold of 5% or less of significant t-tests was used to indicate unidimensionality. The authors hypothesized that all the scales would be unidimensional measuring but one latent trait. Person Separation Index (PSI) was used to assess person reliability.  $PSI \geq 0.80$  was considered acceptable [49,50]. PSI values under 0.80 means that the scale would not be able to distinguish between patients who have high or low foot and

ankle function. Cronbach's alpha was used to assess the internal consistency of the scale. Alpha values  $\geq 0.70$  were considered as acceptable [13].

## 3. Results

Altogether 205 patients with foot and ankle specific problems fulfilled the questionnaires. After excluding the patients with missing values, 183 were included in the analysis. Patients' characteristics are presented in Table 1.

### 3.1. Walking/standing subscale

The mean (SD) score for walking/standing subscale was 57 (24). All response categories were ordered (Fig. 1). Item 3 "I change the way I walk due to pain in my foot" was found potentially misfitting ( $p < 0.001$ ). Residual correlation ( $> 3$ ) was found between items 2 "Avoid walking distances" and 5 "Have to stop and rest".

No statistically significant difference was noted between MTPJ 1 pathologies and other foot and ankle pathologies ( $P = 0.54$ ). However, the walking/standing subscale did not provide coverage for patient locating in  $-5$  and  $+8$  logits (Fig. 2).

The subscale of walking/standing measures one latent trait (Table 2). The unidimensionality was not confirmed as the percentage of significant t-tests was less than 5%.

### 3.2. Pain subscale

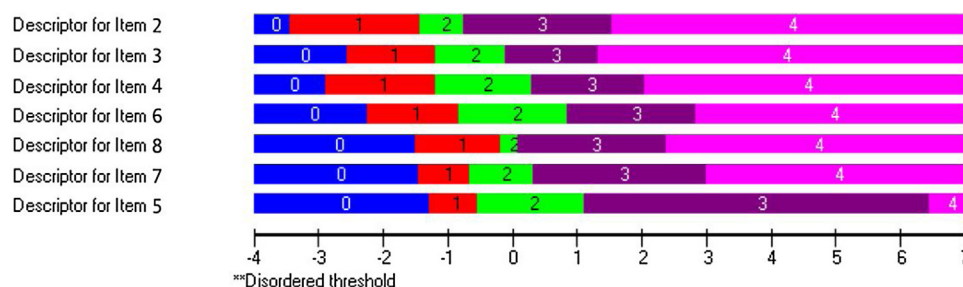
The mean (SD) score for pain subscale was 58 (19). Item 16 showed disordered thresholds in threshold categories (Figs. 3 and 4).

All items showed adequate fit statistics (Table 2). There were no statistically significant probabilities for misfit after Bonferroni correction.

Residual correlation over 0.3 was found between item pairs 1 "Pain in my foot" and 11 "Pain in foot/ankle more painful in evening", 11 and 12 "Shooting pains in foot/ankle", 11 and 15 "Usual pain in foot/ankle", 12 and 16 "Pain from in bed at night" in the refined scale.

The subscale did not provide coverage for two patients with extreme scores in the location of  $-11$  logits in the first run (Fig. 5). Further, the coverage did not exceed to one patient which located in the positive logit six. Nonetheless, the scale had sufficient coverage for the majority of patients with the latent trait of pain.

The unidimensionality of the subscale of pain was not confirmed, as the number of significant t-tests was more than five percent (Table 2). Removing the item 16 led to a unidimensional scale and ordered thresholds in all remaining items (Fig. 6). In the refined scale, two patients still exceeded the range (logit  $-11$  and  $+6$ ) to where the scale provided coverage (Fig. 7). Concerning the distribution of person and item thresholds, there was no



**Fig. 1.** Response categories of the walking/standing subscale. The response categories from left to right are as follows: 0 "None of the time"; 1 "Rarely"; 2 "Some of the time"; 3 "Most of the time"; 4 "All of the time".

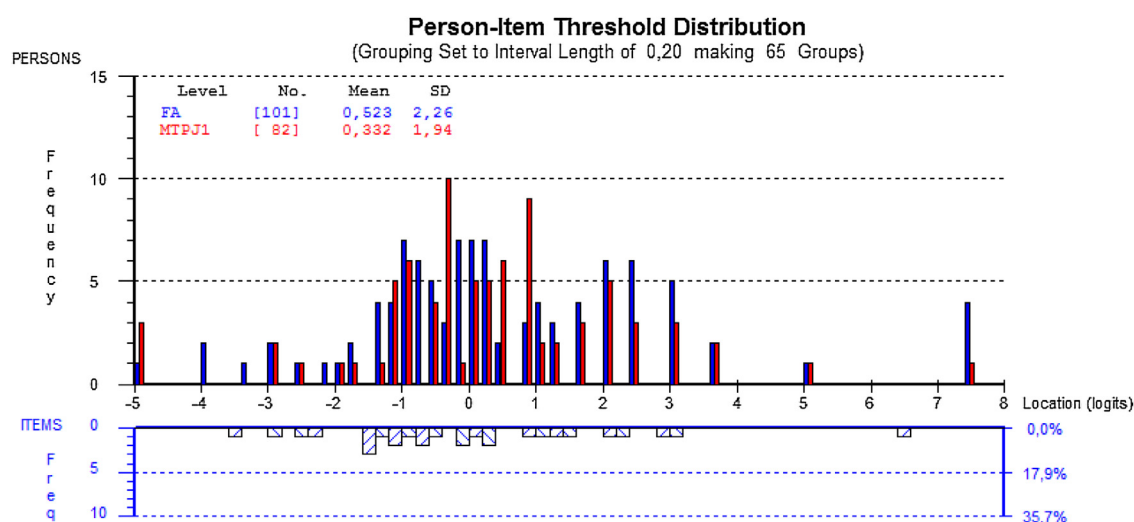


Fig. 2. Person-item threshold distribution of the walking subscale. FA, foot and ankle pathology; MTPJ, first metatarsophalangeal joint pathology.

Table 2  
Analysis specifications.

Subscale	Analysis	Item location	Item fit residual	Person location	Person fit residual	Chi square interaction			Alpha (with extremes)	PSI	% of significant t-tests
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Value	df	P			
Reference values	Initial	0.00 (1.00)						>0.05	>0.7	>0.8	<5.0
Walking/standing	Initial	0.00 (0.80)	0.10 (1.07)	0.44 (2.10)	−0.26 (1.02)	19.24	14	0.16	0.92	0.90	3.83
Pain	Initial	0.00 (1.72)	0.15 (0.58)	0.91 (2.11)	−0.32 (0.99)	6.53	10	0.77	0.83	0.85	6.56
	Remove item 16	0.00 (1.67)	0.16 (0.90)	1.28 (2.16)	−0.35 (0.96)	13.25	8	0.10	0.84	0.84	4.37
Social interaction	Initial	0.00 (0.77)	0.63 (1.57)	0.18 (1.09)	−0.35 (1.06)	20.15	8	<0.01	0.67	0.68	3.80
	Remove item 2	0.00 (1.08)	0.27 (0.60)	1.20 (1.50)	−0.35 (0.96)	4.50	6	0.6	0.70	0.71	0.55

df = degrees of freedom.

P = probability.

PSI = person separation index.

SD = standard deviation.

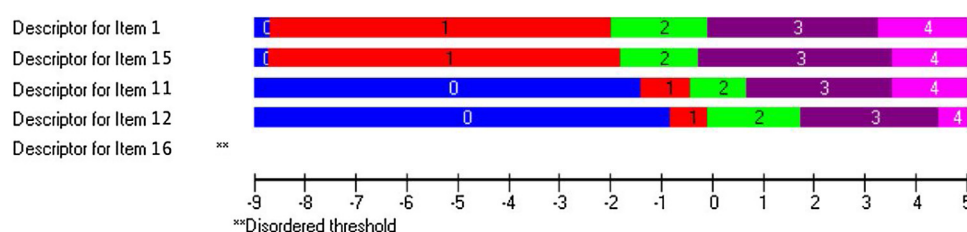


Fig. 3. Response categories of the pain subscale. In items 1, 11 and 12, the response categories from left to right are as follows: 0 “None of the time”; 1 “Rarely”; 2 “Some of the time”; 3 “Most of the time”; 4 “All of the time”. In items 15, the response categories from left to right are: 0 “None”; 1 “Very mild”; 2 “Mild”; 3 “Moderate”; 4 “Severe”. In item 16, the response categories are: 0 “No nights”; 1 “Only 1 or 2 nights”; 2 “Some nights”; 3 “Most nights”; 4 “Every night”.

statistically significant difference in the analyses with the original or the refined scales ( $p = 0.35–0.44$ ) between patients with MTPJ 1 pathology or other foot and ankle regions.

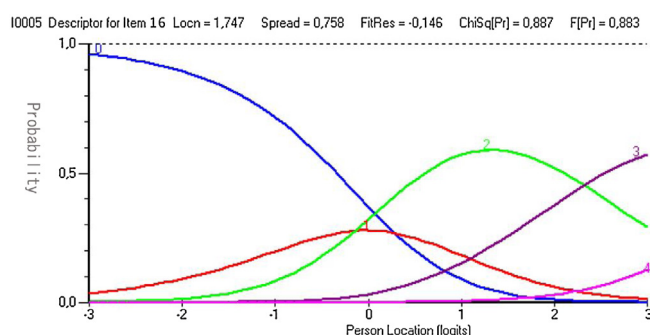
### 3.3. Social interaction subscale

The mean (SD) score for social subscale was 54 (19). Response categories were ordered in items 9 “Feel self-conscious about foot/ankle” and 14 “Unable to do all social or recreational activities” (Fig. 8). However, items 10 “Feel self-conscious about shoes” and 13

“Pain in foot/ankle prevents carrying out work/everyday activities” had disordered thresholds (Fig. 8). Response categories 1 and 2 were collapsed in items 10 and 13. Merging the response categories resulted in ordered thresholds refining the scale (Fig. 9).

The scale did not provide coverage for patients who located on over  $\pm 3$  logits (Fig. 10). However, the scale provided perfect coverage for assessing MTPJ1 pathologies. Nonetheless, there was no statistically significant difference ( $P = 0.46$ ) between the distribution of other patients with foot and ankle pathologies and MTPJ1 pathologies.





**Fig. 4.** Disordered thresholds of item 16 “During the past 4 weeks have you been troubled by pain from your foot in bed at night?” of the pain subscale. The response categories (curves 0–4) of item 16 are with higher scores for pain ascending to the right: 0 “No nights”; 1 “Only 1 or 2 nights”; 2 “Some nights”; 3 “Most nights”; 4 “Every night”. FitRes, Fit residual; F, fit; Locn, location; Pr, probability.

All items had sufficient fit statistics as no Bonferroni-adjusted significance was found (Table 2). Slight residual correlation ( $>0.3$ ) was found between the following items: 9 “I feel self-conscious about my foot” and 10 “I feel self-conscious about the shoes I have to wear”, 10 and 13 “Pain in foot/ankle prevents carrying out work/everyday activities”, 10 and 14 “Unable to do all social or recreational activities”. Mean fit residual indicated potential misfit.

The scale proved to be unidimensional (measure but one trait) as the number of significant paired t-tests was 3.8% (threshold  $<5\%$ ) in equation of item subsets (Table 2).

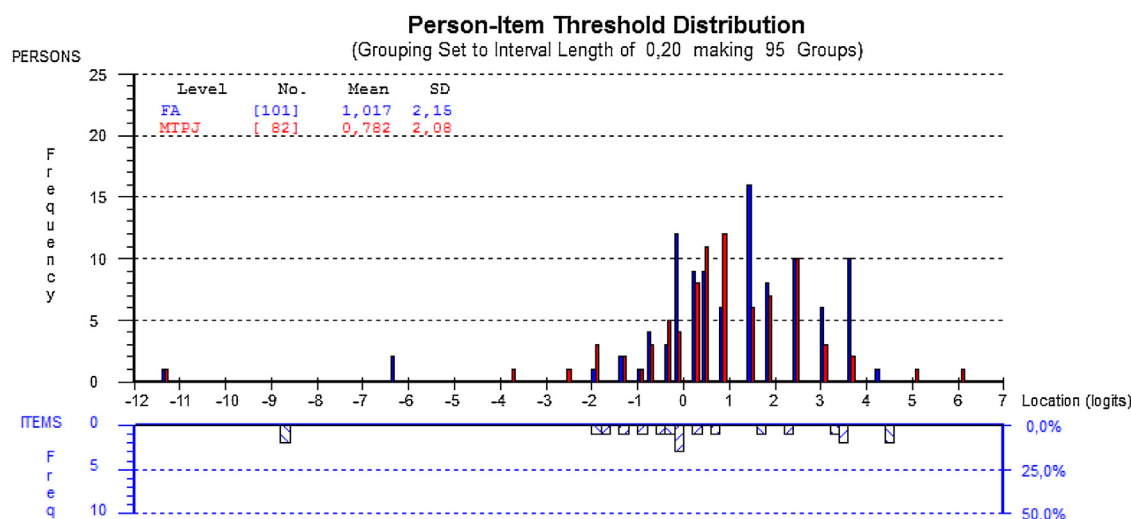
## 4. Discussion

The main result is that the Finnish version of the MOXFQ functions well for orthopaedic foot and ankle patients despite some violations of the Rasch model. The results provide further body of evidence that the MOXFQ has psychometric properties, scientific soundness and relevance suitable for assessing complaints among orthopaedic foot and ankle population.

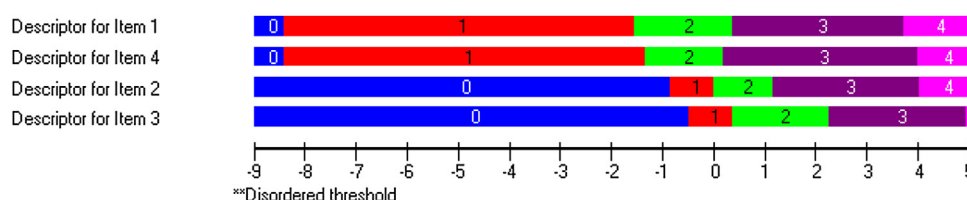
The MOXFQ has been found to fulfill the RMT criteria for hallux valgus patients [16]. In the present study, the item fit of walking/standing and pain subscales was adequate. However, the fit statistics of the social subscale was found to be potentially misfitting. In addition, pain subscale seemed to measure two latent traits but gained unidimensionality when removing item 16 “Have you been troubled by pain from your foot/ankle in bed at night?”. All subscales covered the majority of the patients. However, there were a few patients that exceeded the coverage of the instrument. With testing the removal of item 2 “I avoid walking long distances because of pain in my foot/ankle”, all subscales showed adequate fit statistics. Nonetheless, the scale seems to function quite well with the item 2 included.

The order of the response categories of pain and social interaction was inconsistency. Some refinement could potentially make these response categories function better. These were minor findings but provide useful information for those who use the MOXFQ in foot and ankle patients.

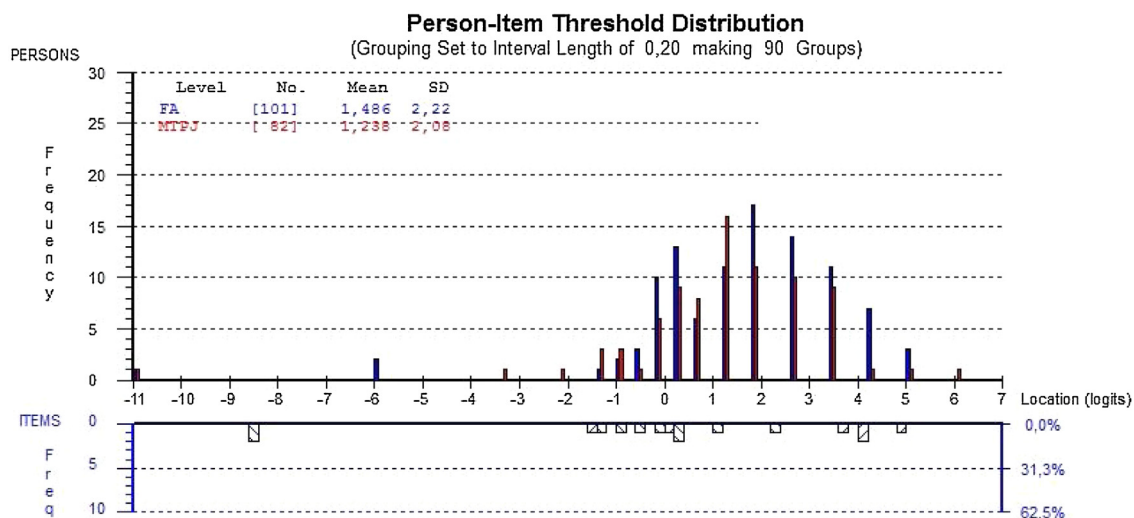
It is important to bear in mind, that the MOXFQ was first developed to measure the outcomes after hallux valgus surgery [16]. During the development process, a set of items were evaluated using the Rasch analysis, and the analysis resulted in a 16-item scale [16]. Jia et al. published a systematic review of



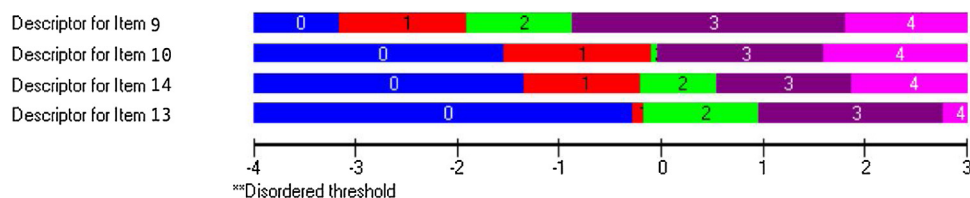
**Fig. 5.** Person-item threshold distribution of the pain subscale. FA, foot and ankle pathology; MTPJ, first metatarsophalangeal joint pathology.



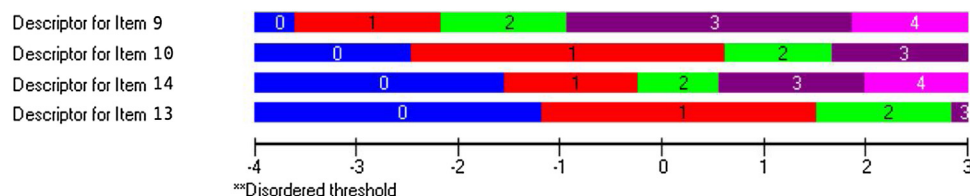
**Fig. 6.** Response categories after removing item 16 in the pain subscale.



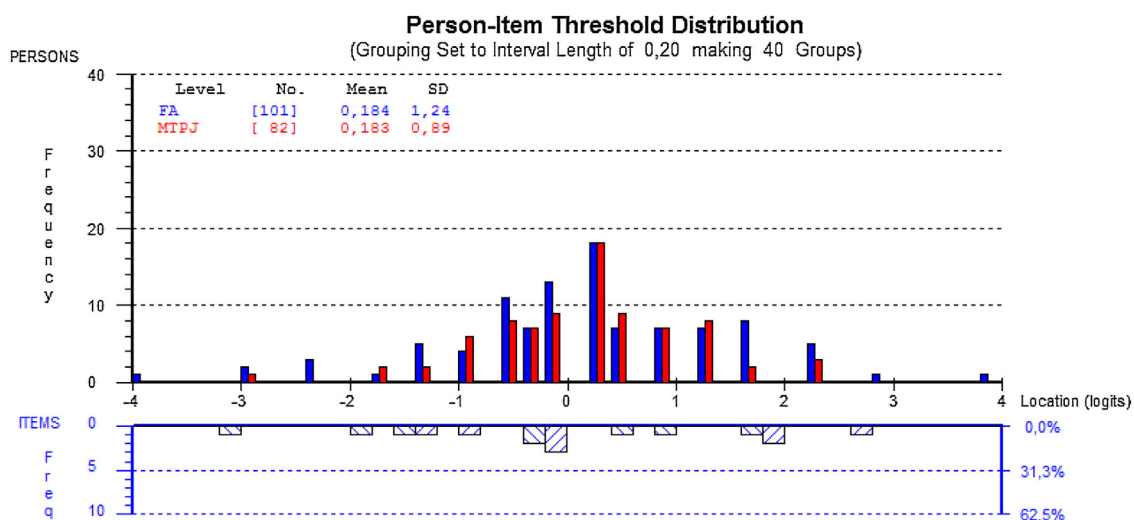
**Fig. 7.** Person-item threshold distribution of the pain subscale after removing item 16. FA, foot and ankle pathology; MTPJ, first metatarsophalangeal joint pathology.



**Fig. 8.** Response categories of the social subscale. The response categories from left to right with 0 "None of the time"; 1 "Rarely"; 2 "Some of the time"; 3 "Most of the time"; 4 "All of the time".



**Fig. 9.** Refined response categories of the Social subscale. Answer categories (1 and 2) of items 10 "Feel self-conscious about shoes" and 13 "Pain in foot/ankle prevents carrying out work/everyday activities" were merged.



**Fig. 10.** Person-item threshold distribution of the Social interaction subscale. FA, foot and ankle pathology; MTPJ, first metatarsophalangeal joint pathology.

psychometric properties of foot and ankle specific PROMs where they concluded that among 50 foot and ankle PROMs, the MOXFQ had superior psychometric properties [8]. The present study found that there were minor violations of the Rasch model when applied to this distinct patient group of orthopaedic foot and ankle patient. However, the problems were somewhat small, and the scale performed well overall. The results support the use of MOXFQ among orthopaedic foot and ankle patients.

The Finnish version of the MOXFQ is available under license agreement from Oxford University Innovation, United Kingdom (<https://innovation.ox.ac.uk/outcome-measures/manchester-oxford-foot-questionnaire-moxfq/>). The MOXFQ can be used to assess patients' status on clinical visits and rehabilitation on follow-ups. The MOXFQ is relatively fast and easy to complete during outpatient visits. The MOXFQ instrument can be used for scientific and research purposes when the patients' voice is needed to be heard in evaluating treatment effectiveness or outcomes. The MOXFQ seems suitable for all foot and ankle patients, but it can especially be used as a region-specific instrument for patients with first metatarsophalangeal joint pathology.

The present study used a sample size adequate for conducting Rasch analysis and this can be mentioned as a strength of the study [34]. Other strengths are the broad spectrum of different foot and ankle problems among this patient group and use of the Rasch analysis. The limitations of this study are its lack of differential item functioning, as the sample size was not large enough for conducting these analyses. Further studies could focus on assessing the differential item functioning between genders and different age groups. Future prospects could also compare the new EFAS score [51] and the MOXFQ in assessing patient-reported outcomes in first metatarsal joint surgery.

In addition, the validation of the MOXFQ for Hindfoot and Ankle surgery could be done over longer timescales to establish its responsiveness to change over a longer period of time.

## 5. Conclusions

Despite some infractions of the Rasch model in the MOXFQ subscale of walking/standing, the instruments functioned well. The subscales of the Finnish MOXFQ are meaningful for assessing patient-reported complaints and outcomes in orthopaedic foot and ankle population.

## Funding

The authors are grateful for financial support from the Musculoskeletal and Plastic Surgery Research Center Helsinki, Helsinki University Hospital and University of Helsinki, the Competitive Research Funding of Helsinki University Hospital and the Finnish Cancer Society.

## Conflict of interest

The authors declared that there is no conflict of interest.

## Acknowledgements

The authors would like to thank Head of Surgery, Dr Maija Pesola for her permission to conduct this study at the Department of Surgery in Central Finland Central Hospital, Jyväskylä, Finland. The authors would also like to thank the nurse staff of Central Finland Central Hospital for their help in the data collection. The authors are grateful to the surgeons who collected data in Central Finland Central Hospital, Peijas Hospital, Seinäjoki Central Hospital and Oulu University Hospital. The authors would like to thank the following persons for help with ensuring that the present study

was successfully completed: Outi Malkavaara, Heli Sarpila, Juhana Leppilähti, Mikko M. Uimonen, Timo Juutilainen, Timo Sirola, Miia Aulanko, Risto Nikku, Risto P. Roine, Anna Sandbacka, Janne Jousmäki, Tomi Tillgren and Mika Sampo.

## References

- [1] Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ* 2002;324(7351):1417.
- [2] Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. *Br Med J* 2015;350:.
- [3] Deshpande PR, Rajan S, Sudeepthi BL, Nazir CA. Patient-reported outcomes: a new era in clinical research. *Perspect Clin Res* 2011;2(4):137.
- [4] Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract* 2006;12(5):559–68.
- [5] Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. 1998.
- [6] Hunt KJMD, Hurwit DBA. Use of patient-reported outcome measures in foot and ankle research. *J Bone Jt Surg Am Vol* 2013;95(16):e118.
- [7] Button G, Pinney S. A meta-analysis of outcome rating scales in foot and ankle surgery: is there a valid, reliable, and responsive system? *Foot Ankle Int* 2004;25(8):521–5.
- [8] Jia Y, Huang H, Gagnier JJ. A systematic review of measurement properties of patient-reported outcome measures for use in patients with foot or ankle diseases. *Qual Life Res* 2017;26(8):1969–2010.
- [9] Van der Leeden M, Steultjens MP, Terwee CB, Rosenbaum D, Turner D, Woodburn J, et al. A systematic review of instruments measuring foot function, foot pain, and foot-related disability in patients with rheumatoid arthritis. *Arthritis Care Res* 2008;59(9):1257–69.
- [10] Eecheute C, Vaes P, Van Aerschoot L, Asman S, Duquet W. The clinimetric qualities of patient-assessed instruments for measuring chronic ankle instability: a systematic review. *BMC Musculoskelet Disord* 2007;8(1):6.
- [11] Shazadeh Safavi P, Janney C, Jupiter D, Kunzler D, Bui R, Panchbhavi VK. A systematic review of the outcome evaluation tools for the foot and ankle. *Foot Ankle Spec* 2018;1938640018803747.
- [12] Terwee CB, Prinsen C, Chiarotto A, de Vet H, Bouter LM, Alonso J, et al. COSMIN methodology for assessing the content validity of PROMs—user manual. 2018.
- [13] Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27(5):1147–57.
- [14] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19(4):539–49.
- [15] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63(7):737–45.
- [16] Dawson J, Coffey J, Doll H, Lavis G, Cooke P, Herron M, et al. A patient-based questionnaire to assess outcomes of foot surgery: validation in the context of surgery for hallux valgus. *Qual Life Res* 2006;15(7):1211–22.
- [17] Schrier JC, Palmieri LN, Verheyen CC, Jansen J, Koeter S. Patient-reported outcome measures in hallux valgus surgery. A review of literature. *Foot Ankle Surg* 2015;21(1):11–5.
- [18] Dawson J, Boller I, Doll H, Lavis G, Sharp R, Cooke P, et al. The MOXFQ patient-reported questionnaire: assessment of data quality, reliability and validity in relation to foot and ankle surgery. *Foot (Edinb)* 2011;21(2):92–102.
- [19] Alvey J, Palmer S, Otter S. A comparison of the readability of two patient-reported outcome measures used to evaluate foot surgery. *J Foot Ankle Surg* 2012;51(4):412–4.
- [20] Venkatesan S, Schotanus MG, Hendrickx RP. Dutch translation of the Manchester-Oxford Foot Questionnaire: reassessment of reliability and validity. *J Foot Ankle Surg* 2016;55(6):1199–201.
- [21] Arbab D, Kuhlmann K, Ringendahl H, Bouillon B, Eysel P, König D. Reliability, validity and responsiveness of the German Manchester-Oxford Foot Questionnaire (MOXFQ) in patients with foot or ankle surgery. *Foot Ankle Surg* 2017.
- [22] Park MJ, Ko YC, Huh JW, Park SH, Park TH, Park JH. Validation of the Korean version of the Manchester-Oxford Foot Questionnaire in patients with hallux valgus. *J Foot Ankle Surg* 2017;56(2):252–4.
- [23] Talu B, Bayramlar K, Bek N, Yakut Y. Validity and reliability of the Turkish version of the Manchester-Oxford Foot Questionnaire for hallux valgus deformity evaluation. *Acta Orthop Traumatol Turc* 2016;50(2):207–13.
- [24] Mousavian A, Ebrahimzadeh MH, Birjandinejad A, Omid-Kashani F, Kachooei AR. Translation and cultural adaptation of the Manchester-Oxford Foot Questionnaire (MOXFQ) into Persian language. *Foot (Edinb)* 2015;25(4):224–7.
- [25] Marinozzi A, Martinelli N, Panasci M, Cancellieri F, Franceschetti E, Vincenzi B, et al. Italian translation of the Manchester-Oxford Foot Questionnaire, with reassessment of reliability and validity. *Qual Life Res* 2009;18(7):923–7.
- [26] Garces JB, Winsor I, Goldhahn S, Castro MD, Swords MP, Grujic L, et al. Reliability, validity and responsiveness of the Spanish Manchester-Oxford Foot Questionnaire (MOXFQ) in patients with foot or ankle surgery. *Foot Ankle Surg* 2016;22(1):59–70.



- [27] Dawson J, Boller I, Doll H, Lavis G, Sharp R, Cooke P, et al. Responsiveness of the Manchester-Oxford Foot Questionnaire (MOXFQ) compared with AOFAS, SF-36 and EQ-5D assessments following foot or ankle surgery. *J Bone Jt Surg Br* 2012;94(2):215–21.
- [28] Andrich D. Rating scales and Rasch measurement. *Expert Rev Pharmacoecon Outcomes Res* 2011;11(5):571–85.
- [29] Elhan A, Küçükdeveci A, Tennant A. The Rasch measurement model. *Res Issues Phys Rehabil Med* 2010;89–102.
- [30] Panayides P, Robinson C, Tymms P. The assessment revolution that has passed England by: Rasch measurement. *Br Educ Res J* 2010;36(4):611–26.
- [31] Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. Applying the Rasch model: fundamental measurement in the human sciences. 2nd ed Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers; 2007 xvi, 340–xvi, 340.
- [32] Tesio L. Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. [Erratum appears in *J Rehabil Med*. 2003 Nov;35(6):294]. *J Rehabil Med* 2003;35(3):105–15.
- [33] Tennant A, Conaghan Philip G. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper?. *Arthritis Care Res* 2007;57(8):1358–62.
- [34] Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN risk of Bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;27(5):1171–9.
- [35] World Medical A. World medical association declaration of helsinki: ethical principles for medical research involving human subjects. *JAMA* 2013;310(20):2191–4.
- [36] Dawson J, Boller I, Doll H, Lavis G, Sharp R, Cooke P, et al. The MOXFQ patient-reported questionnaire: assessment of data quality, reliability and validity in relation to foot and ankle surgery. *Foot* 2011;21(2):92–102.
- [37] Morley D, Jenkinson C, Doll H, Lavis G, Sharp R, Cooke P, et al. The Manchester-oxford Foot Questionnaire (MOXFQ): development and validation of a summary index score. *Bone Jt Res* 2013;2(4):66–9.
- [38] Dawson J, Doll H, Coffey J, Jenkinson C, Oxford, Birmingham F, et al. Responsiveness and minimally important change for the Manchester-Oxford foot questionnaire (MOXFQ) compared with AOFAS and SF-36 assessments following surgery for hallux valgus. *Osteoarthritis Cartil* 2007;15(8):918–31.
- [39] Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value Health* 2005;8(2):94–104.
- [40] Rasch G. Probabilistic models for some intelligence and attainment tests. Danish Institute for Education Research, reprinted: MESA Press; 1993.
- [41] Hagquist C, Bruce M, Gustavsson JP. Using the Rasch model in nursing research: an introduction and illustrative example. *Int J Nurs Stud* 2009;46(3):380–93.
- [42] Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol* 2007;6(12):1094–105.
- [43] Andrich D. Rasch models for measurement, vol. 68. Sage; 1988.
- [44] Linacre JM. Optimizing rating scale category effectiveness. *J Appl Meas* 2002;3(1):85–106.
- [45] Linacre J. Sample size and item calibration stability. *Rasch Mes Trans* 1994;7:328.
- [46] Ehlan A, Kucukdeveci A, Tennant A. The Rasch measurement model. In: Franchignoni F, editor. Research issues in physical & rehabilitation medicine. Pavia: Maugeri Foundation; 2010. p. 89–102.
- [47] Andrich D, Pedler P. A law of ordinal random error: the Rasch measurement model and random error distributions of ordinal assessments. *Measurement* 2019;131:771–81.
- [48] Repo J, Tukiainen E, Roine R, Sampo M, Sandelin H, Häkkinen A. Rasch analysis of the lower extremity functional scale for foot and ankle patients. *Disabil Rehabil* 2018;1–7.
- [49] Smith Jr.[117\_TD\$DIFF] EV. Evidence for the reliability of measures and validity of measure interpretation: a Rasch measurement perspective. *J Appl Meas* 2001.
- [50] Wright BD, Masters GN. Rating scale analysis. Mesa Press; 1982.
- [51] Richter M, Agren P-H, Besse J-L, Cöster M, Kofoed H, Maffulli N, et al. EFAS score—multilingual development and validation of a patient-reported outcome measure (PROM) by the score committee of the European Foot and Ankle Society (EFAS). *Foot Ankle Surg* 2018;24(3):185–204.